FAST VARIATIONAL BAYESIAN SIGNAL RECOVERY IN THE PRESENCE OF POISSON-GAUSSIAN NOISE

Yosra Marnissi[1], *Yuling Zheng*[2], *and Jean-Christophe Pesquet*[1]

[1] Université Paris-Est, LIGM, UMR CNRS 8049, Marne-la-Vallée, France
[2] IBM Research, China

ABSTRACT

This paper presents a new method for solving linear inverse problems where the observations are corrupted with a mixed Poisson-Gaussian noise. To generate a reliable solution, a regularized approach is often adopted in the literature. In this context, the optimal selection of the regularization parameters is of crucial importance in terms of estimation performance. The variational Bayesian-based approach we propose in this work allows us to automatically estimate the original signal and the associated regularization parameter from the observed data. A majorization-minimization technique is employed to circumvent the difficulties raised by the intricate form of the Poisson-Gaussian likelihood. Experimental results show that the proposed method is fast and achieves state-of-the art performance in comparison with approaches where the regularization parameters are manually adjusted.

Index Terms— inverse problems, restoration, reconstruction, Bayesian methods, parameter estimation.

1. INTRODUCTION

During their acquisition process, signals are degraded in various ways - they may be distorted by linear or nonlinear operators and they may be perturbed by different types of noise depending on the system used to capture them. In particular, most imaging systems are photon counting devices, like CMOS and CCD cameras. If such devices are used to acquire an image, the resulting noise is non-additive and signal-dependent and it can be modeled by a Poisson probabilistic distribution. A more realistic approach however consists of also taking into account signal-independent noise sources, often considered as Gaussian, by modelling the noise with a mixed Poisson-Gaussian (PG) distribution. Here, we focus on the recovery of a signal degraded by a linear operator and such a mixed PG noise.

Most existing recovery methods for such data are based on energy minimization approaches. These methods involve the minimization of a well-defined data fidelity term the role of which is to make the solution consistent with the observed data, well-balanced with regularization terms which incorporate a priori information and ensure the stability of the solution. An early work on this type of restoration techniques for images degaded with mixed noise is [26], while more recent developments are presented in [5, 7, 13–15, 22, 23].

The quality of the signals recovered by energy minimization based methods is highly sensitive to the choice of regularization parameters which control the relative weights of the data fidelity and the regularization terms. Too small values of these parameters lead to noisy estimates, while too large values yield oversmoothed estimates. Consequently, an important issue in these methods is the choice of the regularization parameters. If ground truth is available, the regularization parameters can be selected so as to maximize a chosen quality measure, e.g. the Signal-to-Noise Ratio (SNR). In real life applications where ground truth is not available, the selection of the optimal regularization parameters is a more challenging task. Some of the existing methods for hyperparameter estimation are [1,8,10–12,20,21]. Most of the mentioned methods were developed under the assumption that the signals are corrupted with signal independent Gaussian noise, so they are not suitable for optimal parameter determination in inverse problems involving PG noise. One can however mention the works in [16, 17] proposing efficient estimators in the context where no linear degradation operator is present.

Another powerful way of dealing with the regularization parameter choice is to work in the Bayesian framework. In the PG case, Bayesian estimation methods based on MCMC [2] have been recently proposed. However, to our knowledge, no work has been undertaken with variational Bayesian approximation (VBA) techniques [3, 6, 9, 25, 28] which can lead to faster estimation approaches. In this paper, we present a fast VBA approach for signals degraded by an arbitrary linear operator in the presence of a mixed PG noise. Our method aims at providing a good approximate Minimum Mean Square Estimator (MMSE) in the problem of interest. While using VBA, the main difficulty arising in the PG case is that the involved likelihood has a complicated form. To solve this problem, we adopt the Generalized Anscombe Transform (GAST) approximation [18, 27] of the exact PG likelihood. Moreover, a majorization-minimization (MM) technique is adopted providing a tractable VBA solution for nonconjugate distributions [24]. Thanks to the MM strategy, the proposed method allows us to employ a wide class of a priori distributions accounting for the sparsity of the target signal after some appropriate linear transformation.

This paper is organized as follows. In Section 2, we formulate the signal recovery problem for mixed PG noise in the Bayesian framework. In Section 3, we present our proposed estimation method based on VBA. In Section 4 we provide simulation results together with comparisons with state-of-the-art methods in terms of image restoration performance and computation time. Finally, some conclusions are drawn in Section 5.

2. STATEMENT OF THE PROBLEM

We are interested in recovering a signal $\mathbf{x} \in \mathbb{R}^N$ degraded by a matrix $\mathbf{H} \in \mathbb{R}^{M \times N}$ and corrupted with a Poisson-Gaussian noise. More specifically, the vector of observations $\mathbf{y} = (y_i)_{1 \leq i \leq M} \in \mathbb{R}^M$ is related to $\mathbf{x}$ through the following model :

$$\mathbf{y} = \mathbf{z} + \mathbf{b}, \tag{1}$$

where $\mathbf{z}$ and $\mathbf{b}$ are supposed to be mutually independent random vectors and

$$\mathbf{z} \mid \mathbf{x} \sim \mathcal{P}(\mathbf{H}\mathbf{x}), \quad \mathbf{b} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_M),$$

where $\mathcal{P}$ denotes the independent Poisson distribution, $\mathcal{N}$ denotes the Gaussian one, and $\mathbf{I}_M$ is the identity matrix of dimension $M \times M$. Thus, the likelihood function is given by [7]

$$\mathsf{p}(\mathbf{y} \mid \mathbf{x}) = \prod_{i=1}^{M} \left(\sum_{n=1}^{+\infty} \frac{e^{-[\mathbf{H}\mathbf{x}]_i} \left([\mathbf{H}\mathbf{x}]_i\right)^n}{n!} \frac{e^{-\frac{1}{2\sigma^2}(y_i - n)^2}}{\sqrt{2\pi\sigma^2}} \right), \tag{2}$$

where, for every $i \in \{1, \ldots, M\}$, $[\mathbf{H}\mathbf{x}]_i$ denotes the i-th component of $\mathbf{H}\mathbf{x}$. Moreover, we assume the following flexible form of the prior on $\mathbf{x}$:

$$\mathsf{p}(\mathbf{x} \mid \gamma) = \tau \gamma^{\frac{N}{2\kappa}} \exp\Big(-\gamma \sum_{j=1}^{Q} \|\mathbf{D}_j \mathbf{x}\|^{2\kappa} \Big), \tag{3}$$

where κ is a constant in $(0, 1]$, $\|\cdot\|$ denotes the ℓ^2-norm and $\mathbf{D}$ is an operator that can be a matrix computing the discrete difference between neighboring pixels or a frame-analysis operator, which have been blockwise decomposed as $\mathbf{D} = [\mathbf{D}_1^\top, \ldots, \mathbf{D}_Q^\top]^\top$, where each block $(\mathbf{D}_j)_{1 \leqslant j \leqslant Q} \in \mathbb{R}^{S \times N}$. The constant $\gamma \in (0, +\infty)$ can be viewed as a regularization parameter and $\tau \in (0, +\infty)$ is a constant independent of γ. Note that the form of the partition function for such a prior distribution follows from the fact that the associated potential is 2κ-homogeneous [19].

One of the main advantages of our proposed approach is that it allows the parameter γ to be estimated together with $\mathbf{x}$. To this end, we choose a Gamma prior for γ, i.e $\mathsf{p}(\gamma) \propto \gamma^{\alpha-1} \exp(-\beta\gamma)$ where α and β are positive constants.

Using the Bayes' rule, we can obtain the posterior distribution of the set of unknown variables $\mathbf{\Theta} = (\mathbf{x}, \gamma)$, which has an intricate form. In particular, its normalization constant does not have a closed form expression. To cope with this problem, we resort to VBA framework which generates a separable approximation to the original posterior distribution.

3. PROPOSED METHOD

The objective of VBA is to find a separable approximation $q(\mathbf{\Theta}) = \prod_{j=1}^{J} q(\Theta_j)$, which is as close as possible to the posterior distribution $\mathsf{p}(\mathbf{\Theta} \mid \mathbf{y})$, by minimizing the Kullback-Leibler divergence between them :

$$q^{\text{opt}} = \underset{q}{\operatorname{argmin}} \;\; \mathcal{KL}(q(\mathbf{\Theta}) \| \mathsf{p}(\mathbf{\Theta} \mid \mathbf{y})), \tag{4}$$

where

$$\mathcal{KL}(q(\mathbf{\Theta}) \| \mathsf{p}(\mathbf{\Theta} \mid \mathbf{y})) = \int q(\mathbf{\Theta}) \ln \frac{q(\mathbf{\Theta})}{\mathsf{p}(\mathbf{\Theta} \mid \mathbf{y})} \mathrm{d}\mathbf{\Theta}. \tag{5}$$

There exists an optimal solution to the optimization problem in (4), given by

$$(\forall j \in \{1, \ldots, J\}) \;\; q(\Theta_j) \propto \exp\left(\langle \ln \mathsf{p}(\mathbf{y}, \Theta) \rangle_{\prod_{i \neq j} q(\Theta_i)} \right) \tag{6}$$

where $\langle \cdot \rangle_{\prod_{i \neq j} q(\Theta_i)} = \int \cdot \prod_{i \neq j} q(\Theta_i) \mathrm{d}\Theta_i$. Due to the implicit relations existing between $(q(\Theta_j))_{1 \leqslant j \leqslant J}$, an analytical expression of $q(\Theta)$ does not exist. Usually, these distributions are determined in an iterative way, by updating one of the separable components $(q(\Theta_j))_{1 \leqslant j \leqslant J}$ while fixing the others [25].

In this work, we consider the following separable assumption :

$$q(\mathbf{\Theta}) = q(\mathbf{x}) q(\gamma). \tag{7}$$

Unfortunately, by using directly (6), we cannot obtain an explicit expression of $q(\mathbf{x})$ due to the complicated form of the PG likelihood in (2) and of the prior distribution (3). In this paper, we propose to tackle this problem as follows : first, to sidestep the difficulty caused by the PG likelihood, we make use of the Generalized Anscombe Transform (GAST) approximation [18, 27] to the PG followed by a minorization step described below, secondly, for prior density (3), a minorization step is also used which will be explained later on.

Using the GAST approximation, the likelihood of vector $\tilde{\mathbf{y}} \in \mathbb{R}^M$ with components $(\tilde{y}_i)_{1 \leq i \leq M} = 2\big(\sqrt{y_i + \delta}\big)_{1 \leq i \leq M}$, where $\delta = \frac{3}{8} + \sigma^2$, is approximately given by

$$\mathsf{p}(\tilde{\mathbf{y}} \mid \mathbf{x}) = \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\tilde{y}_i - 2\sqrt{[\mathbf{H}\mathbf{x}]_i + \delta} \right)^2 \right). \tag{8}$$

Since the concave function $t \longmapsto \sqrt{t + \delta}$ has a Lipschitz continuous gradient, by introducing an auxiliary vector of variables $\mathbf{w} = (w_i)_{1 \leq i \leq M} \in [0, +\infty)^M$, we can construct a majorant function of the neg-log-likelihood defined as $T(\tilde{\mathbf{y}}, \mathbf{x}; \mathbf{w}) = \sum_{i=1}^{M} T_i(\tilde{y}_i, [\mathbf{H}\mathbf{x}]_i; w_i)$ where, for every $i \in \{1, \ldots, M\}$,

$$\begin{aligned} T_i(\tilde{y}_i, [\mathbf{H}\mathbf{x}]_i; w_i) = &\frac{1}{2}\tilde{y}_i^2 + 2\left([\mathbf{H}\mathbf{x}]_i + \delta\right) - 2\tilde{y}_i\sqrt{w_i + \delta} \\ &- \tilde{y}_i (w_i + \delta)^{-\frac{1}{2}} \left([\mathbf{H}\mathbf{x}]_i - w_i\right) \\ &+ \frac{1}{4}\delta^{-\frac{3}{2}} \tilde{y}_i \left([\mathbf{H}\mathbf{x}]_i - w_i\right)^2. \end{aligned} \tag{9}$$

Similarly, we construct a surrogate function for the prior distribution. Following the same idea as in [6], we obtain the following majorant function for the negative logarithm of the prior distribution :

$$\gamma \sum_{j=1}^{Q} \|\mathbf{D}_j \mathbf{x}\|^{2\kappa} \leqslant \gamma \sum_{j=1}^{Q} \frac{\kappa \|\mathbf{D}_j \mathbf{x}\|^2 + (1-\kappa)\lambda_j}{\lambda_j^{1-\kappa}}. \tag{10}$$

where $(\lambda_j)_{1 \leqslant j \leqslant Q}$ are positive variables. The function in the right-hand side of the above inequality is denoted by $Q(\mathbf{x}, \gamma; \boldsymbol{\lambda}) = \sum_{j=1}^{Q} Q_j(\mathbf{D}_j \mathbf{x}, \gamma; \lambda_j)$ where

$$(\forall j \in \{1, \ldots, Q\}) \quad Q_j(\mathbf{D}_j \mathbf{x}, \gamma; \lambda_j) = \gamma \frac{\kappa \|\mathbf{D}_j \mathbf{x}\|^2 + (1-\kappa)\lambda_j}{\lambda_j^{1-\kappa}}. \tag{11}$$

Then, we can derive the following lower bound of the posterior distribution :

$$\mathsf{p}(\boldsymbol{\Theta} \mid \tilde{\mathbf{y}}) \geqslant L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda}), \tag{12}$$

where function L is defined as

$$L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda}) = C(\tilde{\mathbf{y}}) \exp\left[-T(\tilde{\mathbf{y}},\mathbf{x};\mathbf{w}) - Q(\mathbf{x},\gamma;\boldsymbol{\lambda})\right] \mathsf{p}(\gamma)$$

with $C(\tilde{\mathbf{y}}) = \mathsf{p}(\tilde{\mathbf{y}})^{-1}(2\pi)^{-M/2}\tau\gamma^{\frac{N}{2\kappa}}$. The minorization of the distributions leads to an upper bound of the $\mathcal{KL}$ divergence :

$$\mathcal{KL}(q(\boldsymbol{\Theta})\|\mathsf{p}(\boldsymbol{\Theta} \mid \tilde{\mathbf{y}})) \leqslant \mathcal{KL}(q(\boldsymbol{\Theta})\|L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda})). \tag{13}$$

Thus, Problem (4) can be solved by alternating the following four steps :

- Mimimizing the upper bound in (13) w.r.t. $q(\mathbf{x})$,
- Updating the auxiliary variable $\mathbf{w}$ in order to minimize $\mathcal{KL}(q(\boldsymbol{\Theta})\|L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda}))$,
- Updating the auxiliary variable $\boldsymbol{\lambda}$ in order to minimize $\mathcal{KL}(q(\boldsymbol{\Theta})\|L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda}))$,
- Mimimizing the upper bound in (13) w.r.t. $q(\gamma)$.

The optimizations with respect to $\mathbf{w}$ and $\boldsymbol{\lambda}$ aim at making the bound in (13) as tight as possible. Thanks to the majorization strategy, we can prove that the optimal approximate posterior distribution for $\mathbf{x}$ belongs to the Gaussian family and the optimal approximate posterior distribution for γ belongs to a Gamma one, i.e.

$$q(\mathbf{x}) = \mathcal{N}(\mathbf{m},\boldsymbol{\Sigma}), \quad q(\gamma) = \mathcal{G}(a,b).$$

Therefore, the distribution updates can be performed by updating their parameters, i.e. $\mathbf{m}$, $\boldsymbol{\Sigma}$, a, and b. Subsequently, at a given iteration k of the proposed algorithm, the corresponding estimated variables will be indexed by k.

3.1. Updating $q(\mathbf{x})$

Because of the majorization step, we need to minimize the upper bound on $\mathcal{KL}$ divergence. The standard solution (6) can still be used by replacing the joint distribution by its lower bound $L(\boldsymbol{\Theta},\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda}) \propto L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda})$. As a result, it can be shown that $q^{k+1}(\mathbf{x})$ can be identified as a multivariate Gaussian distribution whose covariance matrix and mean parameter are given by

$$\boldsymbol{\Sigma}_{k+1}^{-1} = \frac{1}{2}\delta^{-\frac{3}{2}}\mathbf{H}^\top \mathrm{Diag}(\tilde{\mathbf{y}})\mathbf{H} + 2\frac{a_k}{b_k}\mathbf{D}^\top\boldsymbol{\Lambda}^k\mathbf{D}, \tag{14}$$

$$\mathbf{m}_{k+1} = \boldsymbol{\Sigma}_{k+1}\mathbf{H}^\top\mathbf{u}, \tag{15}$$

where $\mathbf{u}$ is a $M \times 1$ vector whose i-th component is given by $u_i = \tilde{y}_i(w_i^k+\delta)^{-\frac{1}{2}} + \frac{1}{2}\tilde{y}_i\delta^{-\frac{3}{2}}w_i^k - 2$ and $\boldsymbol{\Lambda}$ is the diagonal matrix whose diagonal elements are $\left(\kappa(\lambda_j^k)^{\kappa-1}\mathbf{I}_S\right)_{1\leq j\leq Q}$. Note that, in high dimension, the vector $\mathbf{m}_{k+1}$ can be computed efficiently by making use of a linear conjugate gradient method.

3.2. Updating $\mathbf{w}$

The auxiliary variable $\mathbf{w}$ is determined by minimizing the upper bound of $\mathcal{KL}$ divergence with respect to this variable :

$$\begin{aligned}\mathbf{w}^{k+1} &= \underset{\mathbf{w}}{\operatorname{argmin}} \int q^{k+1}(\mathbf{x})q^k(\gamma)\ln\frac{q^{k+1}(\mathbf{x})q^k(\gamma)}{L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w},\boldsymbol{\lambda}^k)}\mathrm{d}\mathbf{x}\mathrm{d}\gamma \\ &= \underset{\mathbf{w}}{\operatorname{argmin}} \sum_{i=1}^{M} T_i(\tilde{y}_i,[\mathbf{H}\mathbf{m}_{k+1}]_i;w_i), \end{aligned} \tag{16}$$

where the last equality follows from the expression in (9). Actually, $T_i(\tilde{y}_i,[\mathbf{H}\mathbf{m}_{k+1}]_i;w_i)$ is differentiable with respect to w_i and there exist two critical points : $w_i = 0$ and $w_i = [\mathbf{H}\mathbf{m}_{k+1}]_i$. The minimum is attained at

$$w_i^{k+1} = \max\{[\mathbf{H}\mathbf{m}_{k+1}]_i, 0\}. \tag{17}$$

3.3. Updating $\boldsymbol{\lambda}$

The variable $\boldsymbol{\lambda}$ is determined in a similar way :

$$\lambda_j^{k+1} = \underset{\lambda_j\in[0,+\infty)}{\operatorname{argmin}} \; \mathcal{KL}(q^{k+1}(\mathbf{x})q^k(\gamma)\|L(\boldsymbol{\Theta}|\tilde{\mathbf{y}};\mathbf{w}^{k+1},\boldsymbol{\lambda})). \tag{18}$$

Again, $L(\boldsymbol{\Theta},\tilde{\mathbf{y}};\mathbf{w}^{k+1},\boldsymbol{\lambda})$ is differentiable w.r.t. $(\lambda_j)_{1\leqslant j\leqslant N}$. The minimum is attained at

$$\begin{aligned}\lambda_j^{k+1} &= \mathbb{E}_{q^{k+1}(\mathbf{x})}\left[\|\mathbf{D}_j\mathbf{x}\|^2\right] \\ &= \|\mathbf{D}_j\mathbf{m}_{k+1}\|^2 + \mathrm{trace}\left[\mathbf{D}_j^\top\mathbf{D}_j\boldsymbol{\Sigma}_{k+1}\right]. \end{aligned} \tag{19}$$

3.4. Updating $q(\gamma)$

Using (6) where the joint distribution is replaced by its lower bound function, we obtain

$$\begin{aligned} q^{k+1}(\gamma) \propto\, &\gamma^{\frac{N}{2\kappa}+\alpha-1}\exp(-\beta\gamma) \\ &\times\exp\left(-\gamma\sum_{j=1}^{Q}\frac{\kappa\mathbb{E}_{q^{k+1}(\mathbf{x})}\left[\|\mathbf{D}_j\mathbf{x}\|^2\right]+(1-\kappa)\lambda_j^{k+1}}{(\lambda_j^{k+1})^{1-\kappa}}\right). \end{aligned} \tag{20}$$

Using (19), one can recognize that the above distribution is a Gamma one with parameters

$$a_{k+1} = \frac{N}{2\kappa} + \alpha = a, \qquad b_{k+1} = \sum_{j=1}^{Q}(\lambda_j^{k+1})^\kappa + \beta. \tag{21}$$

Finally, our algorithm can be summed up as follows :

1. Set initial values : $\mathbf{w}^0, \boldsymbol{\lambda}^0, b_0$. Compute a with (21).
2. For $k = 0, 1, \ldots$
   (a) Update parameters $\boldsymbol{\Sigma}_{k+1}$ and $\mathbf{m}_{k+1}$ of $q^{k+1}(\mathbf{x})$ using (14) and (15).
   (b) Update $\mathbf{w}^{k+1}$ using (17).
   (c) Update $\boldsymbol{\lambda}^{k+1}$ using (19).
   (d) Update parameter b_{k+1} of $q^{k+1}(\gamma)$ using (21).

4. SIMULATION RESULTS

We evaluate the performance of the proposed approach for the restoration of images degraded by both blur and Poisson-Gaussian noise. In our experiments, we choose a total variation prior, i.e. $\kappa = 1/2$ and for every pixel $j \in \{1,\ldots,N\}$, $\mathbf{D}_j\mathbf{x} = \left[[\nabla^h\mathbf{x}]_j, [\nabla^v\mathbf{x}]_j\right]^\top \in \mathbb{R}^2$ where ∇^h and ∇^v are the discrete gradients computed in the horizontal and vertical directions. As a result, $Q = N$.

		MAP (GAST) [4]	MAP (GAST) [7]	MAP(EXP)	MAP (Exact)	BV (GAST)
First image (350×350) : $x^+ = 20$	γ	fixed	fixed	fixed	fixed	automatic
h : Uniform 5×5, $\sigma^2 = 9$	SNR	13.61	13.60	13.72	13.73	13.80
	Time (s.)	2897	490	3124	48587	29
Second image (257×256) : $x^+ = 60$	γ	fixed	fixed	fixed	fixed	automatic
h : Gaussian 9×9, std 0.5, $\sigma^2 = 36$	SNR	15.35	15.33	15.42	15.43	15.22
	Time (s.)	3168	86	112	612	7
Third image (256×256) : $x^+ = 100$	γ	fixed	fixed	fixed	fixed	automatic
h : Uniform 3×3, $\sigma^2 = 36$	SNR	13.71	13.77	13.81	13.81	14.17
	Time (s.)	2921	578	1060	17027	9
Fourth image (256×256) : $x^+ = 150$	γ	fixed	fixed	fixed	fixed	automatic
h : Gaussian 7×7, std 1, $\sigma^2 = 40$	SNR	20.17	20.11	20.11	20.33	20.43
	Time (s.)	2964	886	3026	43397	14

Table 1: Restoration results for four test images : First image (initial SNR=7.64 dB), second image (initial SNR=9.40 dB), third image (initial SNR=10.68 dB), fourth image (initial SNR= 15.77 dB).

We consider four test images whose intensities have been rescaled so that pixel values belong to a chosen interval $[0, x^+]$. These images are artificially degraded by an operator $\mathbf{H}$ modeling spatially invariant blur with point spread function h and by Poisson noise. Moreover, we add zero-mean white Gaussian noise with variance σ^2. Figure 1 allows us to evaluate the resulting visual improvement for the reconstructed images. The gain is confirmed by the perceptual measures of visual similarity (SSIM) indicated in the figure caption.

The proposed algorithm is compared with recent methods. First, we consider the method described in [4] for recovering images degraded with blur and mixed noise, based on GAST approximation using a smoothed version of the total variation regularization and a spectral projected gradient method for minimizing the energy function. Second, we use the MAP estimate presented in [7] with the non-smoothed form of the total variation under the GAST approximation, the Exponential (EXP) approximation [5], and the exact likelihood (2), where a primal-dual splitting algorithm is adopted to minimize the associated penalized criterion. For those approaches, the regularization parameter γ is adjusted empirically to achieve the maximum SNR value.

Table 1 reports the results obtained with the different algorithms in terms of SNR and approximate computation time needed for convergence. Simulations are performed on an Intel(R) Xeon(R) CPU E5-2630, @ 2.40 GHz, using a Matlab 7 implementation and by taking as initial value the degraded image. We can observe that these approaches yield comparable quantitative results. It is worth noticing that the problem of setting the regularization parameter for these algorithms must be carefully addressed as it highly impacts the quality of the restored image. The main advantage of our approach is that this parameter is tuned automatically, while the proposed algorithm also appears to be the most competitive in terms of computation time.

5. CONCLUSION

In this paper, we have proposed a variational Bayesian method to recover signals degraded with a blur and a mixed Poisson-Gaussian noise, using a GAST approximation of the likelihood. Instead of setting it empirically, the regularization parameter is tuned automatically during the recovery process. Experiments carried out on several images have shown that the proposed approach achieves a good tradeoff between estimation performance and computation time. Several ideas could be investigated in our future work such as the ability of the proposed approach to be extended to more general prior distributions than the one presented in this paper and to other approximations of the Poisson-Gaussian data fidelity term.

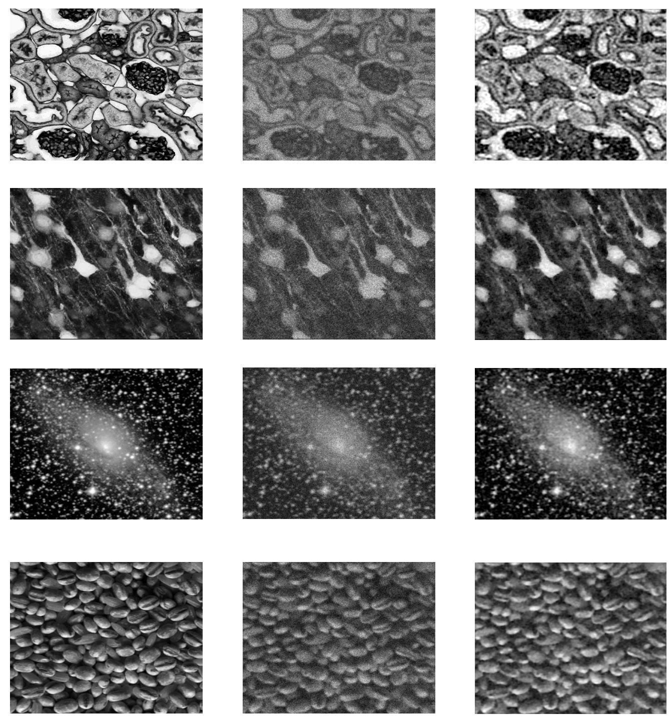

Fig. 1: From left to right : Original images. Degraded images : SSIM=(0.232, 0.423, 0.561, 0.586). Restored images with the proposed approach : SSIM=(0.575, 0.653, 0.765, 0.830).

Acknowlegments : We would like to thank Buda Bajić from Novi Sad University for her help in writing this paper and for running her codes [4] on our image restoration examples.

6. REFERENCES

[1] M. S. C. Almeida and M. Figueiredo, "Parameter estimation for blind and non-blind deblurring using residual whiteness measures," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2751–2763, 2013.

[2] Y. Altmann, A. Wallace, and S. McLaughlin, "Spectral unmixing of multispectral Lidar signals," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5525–5533, Oct. 2015.

[3] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian super resolution," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 984–999, 2011.

[4] B. Bajić, J. Lindblad, and N. Sladoje, "A comparative study of restoration methods for images degraded with Gaussian, Poisson and Poisson-Gaussian noise," *submitted to J. Math. Imaging Vis.*, 2015.

[5] F. Benvenuto, A. La Camera, C. Theys, A. Ferrari, H. Lantéri, and M. Bertero, "The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise," *Inverse Probl.*, vol. 24, no. 3, 2008, 20 pp.

[6] Z. Chen, S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational bayesian methods for multimedia problems," *IEEE Trans. Multimedia*, vol. 16, no. 4, pp. 1000–1017, 2014.

[7] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot, "A convex approach for image restoration with exact Poisson-Gaussian likelihood," *SIAM J. Imag. Sci.*, vol. 8, no. 4, pp. 2662–2682, 2015.

[8] C.-A. Deledalle, S. Vaiter, J. Fadili, and G. Peyré, "Stein Unbiased GrAdient estimator of the risk (SUGAR) for multiple parameter selection," *SIAM J. Imag. Sci.*, vol. 7, no. 4, pp. 2448–2487, 2014.

[9] A. Drémeau, C. Herzet, and L. Daudet, "Boltzmann machine and mean-field approximation for structured sparse decompositions," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3425–3438, 2012.

[10] Y. C. Eldar, "Generalized SURE for exponential families : Applications to regularization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 471–481, 2009.

[11] R. Giryes, M. Elad, and Y. C. Eldar, "The projected GSURE for automatic parameter tuning in iterative shrinkage methods," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 3, pp. 407–422, 2011.

[12] P. C. Hansen, M. E. Kilmer, and R. H. Kjeldsen, "Exploiting residual information in the parameter choice for discrete ill-posed problems," *BIT*, vol. 46, no. 1, pp. 41–59, 2006.

[13] A. Jezierska, E. Chouzenoux, J.-C. Pesquet, and H. Talbot, "A primal-dual proximal splitting approach for restoring data corrupted with Poisson-Gaussian noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Kyoto, Japan, 25-30 Mar. 2012, pp. 1085–1088.

[14] H. Lantéri and C. Theys, "Restoration of astrophysical images - the case of Poisson data with additive Gaussian noise," *EURASIP J. Adv. Signal Process.*, vol. 2005, no. 15, pp. 2500–2513, 2005.

[15] J. Li, Z. Shen, R. Jin, and X. Zhang, "A reweighted ℓ_2 method for image restoration with Poisson and mixed Poisson-Gaussian noise," *Inverse Probl. Imaging*, vol. 9, no. 3, pp. 875–894, Aug. 2015.

[16] F. Luisier, T. Blu, and M. Unser, "Image denoising in mixed Poisson-Gaussian noise," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 696–708, 2011.

[17] M. Mäkitalo and A. Foi, "Poisson-Gaussian denoising using the exact unbiased inverse of the generalized Anscombe transformation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Kyoto, Japan, 25-30 March 2012, pp. 1081–1084.

[18] F. Murtagh, J.-L. Starck, and A. Bijaoui, "Image restoration with noise suppression using a multiresolution support," *Astron. Astrophys. Suppl.*, vol. 112, pp. 179–189, 1995.

[19] M. Pereyra, J. Bioucas-Dias, and M. A. T. Figueiredo, "Maximum-a-posteriori estimation with unknown regularisation parameters," in *Proc. Eur. Sig. Proc. Conf.*, Nice, France, 31 Aug. - 4 Sept. 2015, pp. 2244–2248.

[20] J.-C. Pesquet, A. Benazza-Benyahia, and C. Chaux, "A SURE approach for digital signal/image deconvolution problems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4616–4632, 2009.

[21] S. Ramani, T. Blu, and M. Unser, "Monte-Carlo SURE : A black-box optimization of regularization parameters for general denoising algorithms," *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1540–1554, 2008.

[22] A. Repetti, E. Chouzenoux, and J.-C. Pesquet, "A penalized weighted least squares approach for restoring data corrupted with signal-dependent noise," in *Proc. Eur. Sig. Proc. Conf.*, Bucharest, Roumania, 27-31 Aug. 2012, pp. 1553–1557.

[23] T. D. Roberts and N. G. Kingsbury, "Bayesian denoising/deblurring of Poisson-Gaussian corrupted data using complex wavelets," in *Proc. IEEE Int. Symp. Biomed. Imaging*, Beijing, China, April 2014.

[24] M. Seeger and G. Bouchard, "Fast variational Bayesian inference for non-conjugate matrix factorization models," in *Proc. 15th Int. Conf. Artif. Intell. Stat.*, La Palma, Canary Islands, April 21-23 2012, pp. 1012–1018.

[25] V. Šmídl and A. Quinn, *The Variational Bayes Method in Signal Processing.* New York, USA : Springer-Verlag, 2005.

[26] D. L. Snyder, A. M. Hammoud, and R. L. White, "Image recovery from data acquired with a charge-coupled-device camera," *J. Opt. Soc. Am. A*, vol. 10, no. 5, pp. 1014–1023, May 1993.

[27] J.-L. Starck and F. Murtagh, "Automatic noise estimation from the multiresolution support," *Publications of the Astronomical Society of the Pacific*, vol. 110, no. 744, pp. 193–199, 1998.

[28] Y. Zheng, A. Fraysse, and T. Rodet, "Efficient variational Bayesian approximation method based on subspace optimization," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 681–693, 2015.